

PENGUATAN LITERASI ASESMEN GURU MELALUI PELATIHAN ANALISIS BUTIR SOAL MENGGUNAKAN SOFTWARE 3S-CTT

Johri Sabaryati^{1*}, Linda Sekar Utami², Zulkarnain³, Muhammad Isnaini⁴,
Islahudin⁵, Ahyati Kurniamala Niswariyana⁶

^{1,2,3,4,5}Program Studi Pendidikan Fisika, Universitas Muhammadiyah Mataram, Indonesia

⁶Program Studi Pendidikan Bahasa dan Sastra Indonesia, Universitas Muhammadiyah Mataram, Indonesia

joyafarashy@gmail.com

ABSTRAK

Abstrak: Analisis butir soal merupakan tahapan penting untuk memastikan kualitas instrumen evaluasi, namun praktik ini belum dilakukan secara sistematis oleh sebagian guru karena keterbatasan penguasaan statistik, kesiapan data, dan hambatan penggunaan perangkat lunak. Kegiatan pengabdian kepada masyarakat ini bertujuan memperkuat literasi asesmen guru melalui pelatihan analisis butir soal berbasis *Classical Test Theory* menggunakan software 3S-CTT. Mitra kegiatan adalah SMP Muhammadiyah di Kabupaten Lombok Barat Nusa Tenggara Barat dengan peserta 10 orang guru. Data praktik menggunakan paket soal pilihan ganda sebanyak 30 butir beserta respons siswa atau data latihan yang disiapkan oleh tim. Metode kegiatan disusun dalam tiga tahap, yaitu pra-kegiatan, pelaksanaan, dan evaluasi. Tahap pra-kegiatan meliputi koordinasi mitra, identifikasi kebutuhan, penyusunan modul, dan penyiapan data latihan. Tahap pelaksanaan mencakup sosialisasi konsep CTT, demonstrasi 3S-CTT, praktik input data, analisis, dan interpretasi *output*. Tahap evaluasi dilakukan melalui *pre-test*, *post-test*, observasi, checklist keterampilan, telaah *output*, dan diskusi reflektif. Hasil kegiatan menunjukkan bahwa pelatihan berbasis praktik membantu guru memahami indikator tingkat kesukaran, daya pembeda, reliabilitas, dan efektivitas pengecoh serta menyusun rekomendasi perbaikan butir. Kegiatan ini diharapkan menjadi model penguatan evaluasi pembelajaran berbasis data di sekolah mitra.

Kata Kunci: Analisis Butir; *Classical Test Theory*; Literasi Asesmen; Pelatihan Guru; 3S-CTT.

Abstract: Item analysis is an important stage in ensuring the quality of assessment instruments; however, it has not been implemented systematically by many teachers due to limited statistical knowledge, data readiness, and software-related barriers. This community service program aimed to strengthen teachers' assessment literacy through item analysis training based on *Classical Test Theory* using 3S-CTT software. The partner was a Muhammadiyah of junior secondary school in West Lombok, West Nusa Tenggara, involving 10 teachers. The practice used a 30-item multiple-choice test package along with students' response data or training data prepared by the team. The program was organized into three stages: pre-activity, implementation, and evaluation. The pre-activity stage included partner coordination, needs identification, module development, and preparation of practice data. The implementation stage covered CTT concept introduction, 3S-CTT demonstration, data input practice, analysis, and output interpretation. The evaluation stage used pre-test, post-test, observation, skill checklist, output review, and reflective discussion. The results indicate that practice-based training helped teachers understand item difficulty, discrimination power, reliability, and distractor effectiveness and formulate item improvement recommendations. This program is expected to serve as a model for strengthening data-driven classroom assessment in partner schools.

Keywords: Item Analysis; *Classical Test Theory*; Assessment Literacy; Teacher Training; 3S-CTT.



Article History:

Received: 13-05-2026

Revised : 02-07-2026

Accepted: 03-07-2026

Online : 01-08-2026



This is an open access article under the
CC-BY-SA license

A. LATAR BELAKANG

Kualitas evaluasi pembelajaran tidak hanya ditentukan oleh kesesuaian soal dengan tujuan pembelajaran, tetapi juga oleh mutu butir soal yang digunakan untuk mengukur capaian peserta didik. Butir yang terlalu mudah, terlalu sukar, tidak mampu membedakan kemampuan peserta didik, atau memiliki pengecoh yang tidak berfungsi dapat menurunkan ketepatan informasi hasil belajar (Qadir et al., 2024). Analisis butir diperlukan untuk memperoleh bukti empiris tentang tingkat kesukaran, daya pembeda, reliabilitas, dan efektivitas pengecoh sebagai dasar perbaikan instrumen evaluasi (Ayanwale et al., 2022; Hu et al., 2021; Rezigalla et al., 2024; Vicent & Shanmugam, 2020). Dengan demikian analisis butir soal sangat penting dalam evaluasi pembelajaran terutama analisis menggunakan teori tes klasik atau *Classical Test Theory* (CTT).

Classical Test Theory (CTT) masih relevan digunakan dalam konteks evaluasi pembelajaran di sekolah karena prosedurnya relatif mudah dipahami, hasilnya dapat diinterpretasikan secara langsung, dan indikatornya sesuai dengan kebutuhan guru dalam menilai kualitas soal. Melalui CTT, guru dapat menentukan butir yang layak dipertahankan, direvisi, atau diganti berdasarkan bukti kuantitatif (Ridha et al., 2025). Studi terbaru menunjukkan bahwa analisis berbasis CTT efektif digunakan untuk membaca indeks kesukaran, daya pembeda, efektivitas pengecoh, dan reliabilitas pada tes pilihan ganda, termasuk ketika analisis dilakukan dengan bantuan perangkat lunak seperti AnBuso, Iteman, Klasika, R, maupun perangkat lain yang sejenis (Ayanwale et al., 2022; Hayat, 2021; Nurjanah et al., 2024; Rohmatdi et al., 2024). Dengan demikian CTT dapat membantu guru menganalisis secara praktis butir soal dan praktis, dan indikatornya sesuai dengan kebutuhan guru dalam menilai mutu butir soal.

Meskipun analisis butir penting, praktiknya di sekolah mitra belum berjalan secara berkelanjutan. Berdasarkan identifikasi awal, guru telah terbiasa menyusun soal untuk penilaian harian, penilaian tengah semester, dan penilaian akhir semester, tetapi evaluasi kualitas butir lebih banyak dilakukan melalui telaah kualitatif atau pemeriksaan manual sederhana. Kondisi ini menunjukkan adanya kesenjangan antara kebutuhan peningkatan mutu asesmen dengan kemampuan teknis guru dalam mengolah data respons peserta didik secara kuantitatif.

Beberapa perangkat analisis, seperti Iteman, spreadsheet, R, dan aplikasi psikometri lainnya, dapat digunakan untuk analisis butir. Namun, penggunaannya sering membutuhkan format data tertentu, pengaturan awal yang tepat, serta kemampuan membaca *output statistic* (Susanto et al., 2022). Kajian tentang perangkat analisis butir menunjukkan bahwa pilihan aplikasi perlu mempertimbangkan kompleksitas analisis, ukuran data, kemudahan penggunaan, dan tingkat kemampuan pengguna (Choi & Asilkalkan, 2019; Hayat, 2021; Susanto et al., 2025, 2026). Selain itu, penelitian tentang integrasi teknologi pendidikan menegaskan bahwa

hambatan penggunaan teknologi tidak hanya terkait ketersediaan perangkat, tetapi juga kesiapan, pelatihan, dukungan teknis, dan keyakinan pengguna dalam memanfaatkan teknologi secara bermakna (Alatli, 2020; Atabek, 2020; Gonzales et al., 2025; Shahzad et al., 2025; Theodorio, 2024). Dengan demikian, pemilihan perangkat analisis butir perlu disesuaikan dengan kebutuhan analisis, karakteristik data, kemudahan penggunaan, dan kemampuan pengguna dalam memahami *output statistic*.

Dalam konteks tersebut, software 3S-CTT diperkenalkan sebagai alternatif yang lebih sederhana, ringan, dan ramah pengguna untuk membantu guru melakukan analisis butir berbasis CTT. Penggunaan perangkat lunak yang mudah diakses dan didampingi dengan praktik langsung dipandang relevan untuk mengurangi hambatan teknis guru. Oleh karena itu, kegiatan pengabdian kepada masyarakat ini bertujuan memperkuat literasi asesmen guru melalui pelatihan analisis butir soal menggunakan 3S-CTT. Secara khusus, kegiatan ini diarahkan agar guru mampu menyiapkan data respons, menjalankan analisis, membaca *output* tingkat kesukaran, daya pembeda, reliabilitas, dan efektivitas pengecoh, serta menyusun rekomendasi perbaikan butir berdasarkan hasil analisis.

B. METODE PELAKSANAAN

Kegiatan pengabdian kepada masyarakat ini dilaksanakan pada salah satu sekolah menengah pertama muhammadiyah di Nusa Tenggara Barat dengan sasaran 10 orang guru. Penulisan lokasi dibuat secara umum untuk menjaga fokus artikel pada proses pelatihan dan capaian program. Kegiatan dilaksanakan melalui pendekatan pelatihan dan pendampingan berbasis praktik karena keterampilan analisis butir tidak cukup dibangun melalui ceramah, tetapi memerlukan demonstrasi, latihan langsung, umpan balik, dan interpretasi *output*. Metode pelaksanaan disusun dalam tiga tahap utama, yaitu tahap pra-kegiatan, tahap pelaksanaan, dan tahap evaluasi, dijelaskan secara rinci sebagai berikut:

1. Tahap pra kegiatan difokuskan pada koordinasi dengan mitra, identifikasi kebutuhan guru, penyusunan modul pelatihan, penyiapan akun atau akses software, serta penyiapan data latihan. Data praktik berupa paket soal pilihan ganda sebanyak 30 butir beserta data respons siswa atau data latihan yang disediakan tim.
2. Tahap pelaksanaan meliputi sosialisasi konsep CTT, demonstrasi penggunaan 3S-CTT, praktik input data, pengaturan kunci jawaban, proses analisis, serta pendampingan interpretasi *output*.
3. Tahap evaluasi dilakukan untuk mengukur keberhasilan kegiatan melalui *pre-test*, *post-test*, observasi, checklist keterampilan, telaah *output* peserta, dan diskusi reflektif.

Tahapan pelaksanaan dan teknik evaluasi kegiatan serta indikator capaian dirangkum pada Tabel 1 dan Tabel 2.

Tabel 1. Tahapan Pelaksanaan dan Teknik Evaluasi Kegiatan

Tahap	Fokus Kegiatan	Bentuk Kegiatan	Instrumen/Bukti Evaluasi
Pra-kegiatan	Persiapan program dan kebutuhan mitra	Koordinasi dengan sekolah mitra, identifikasi pengalaman guru dalam menyusun dan menganalisis soal, penyusunan modul, penyiapan 3S-CTT, dan penyiapan data latihan.	Catatan koordinasi, daftar peserta, modul pelatihan, data latihan, dan lembar <i>pre-test</i> .
Pelaksanaan	Pelatihan dan praktik penggunaan 3S-CTT	Sosialisasi konsep CTT, demonstrasi software, praktik input data, pengaturan kunci jawaban, menjalankan analisis, dan membaca <i>output</i> .	Dokumentasi kegiatan, checklist keterampilan, file <i>output</i> analisis, dan observasi proses praktik.
Evaluasi	Pengukuran capaian dan tindak lanjut	<i>Post-test</i> , telaah laporan praktik, diskusi interpretasi hasil, refleksi kendala, serta penyusunan rencana keberlanjutan.	Skor <i>pre-test</i> dan <i>post-test</i> , rubrik interpretasi, catatan diskusi, dan umpan balik peserta.

Evaluasi kegiatan dilakukan secara deskriptif kuantitatif dan kualitatif. Evaluasi kuantitatif menggunakan *pre-test* dan *post-test* untuk melihat perubahan pemahaman dasar peserta tentang analisis butir. Evaluasi keterampilan dilakukan melalui checklist yang menilai kemampuan peserta menyiapkan data, melakukan input, mengatur kunci jawaban, menjalankan analisis, menyimpan *output*, dan membaca hasil. Evaluasi kualitatif dilakukan melalui observasi, telaah laporan praktik, diskusi tanya jawab, dan refleksi kendala yang dialami peserta selama menggunakan 3S-CTT.

Tabel 2. Indikator Capaian Kegiatan

No.	Aspek Capaian	Target/Indikator	Cara Mengukur
1	Pemahaman konsep	Minimal 90% peserta memahami konsep CTT, tingkat kesukaran, daya pembeda, reliabilitas, dan efektivitas pengecoh.	<i>Pre-test</i> dan <i>post-test</i> .
2	Keterampilan teknis	Minimal 85% peserta mampu menyiapkan data, melakukan input, mengatur kunci, dan menjalankan analisis secara mandiri.	Observasi dan checklist keterampilan.
3	<i>Output</i> praktik	Setiap peserta atau kelompok menghasilkan minimal satu file <i>output</i> analisis butir.	File <i>output</i> /laporan praktik.
4	Interpretasi hasil	Minimal 80% peserta mampu menjelaskan hasil analisis dan menyusun rekomendasi perbaikan butir.	Rubrik interpretasi dan diskusi tanya jawab.

C. HASIL DAN PEMBAHASAN

1. Tahap Pra-Kegiatan: Identifikasi Kebutuhan dan Penyiapan Pelatihan

Tahap pra-kegiatan menghasilkan gambaran awal mengenai kebutuhan mitra. Guru telah memiliki pengalaman menyusun soal untuk evaluasi pembelajaran, tetapi sebagian besar belum melakukan analisis butir secara kuantitatif dan berkelanjutan. Analisis kualitas soal masih lebih banyak dilakukan melalui pertimbangan isi, pengalaman mengajar, dan pemeriksaan jawaban siswa secara manual. Kondisi ini menunjukkan bahwa pelatihan tidak hanya perlu mengenalkan konsep analisis butir, tetapi juga menyediakan alur kerja yang sederhana agar guru dapat mengolah data respons siswa secara mandiri.

Hasil identifikasi juga menunjukkan bahwa fasilitas dasar sekolah cukup mendukung, seperti ketersediaan perangkat laptop guru, ruang pelatihan, akses listrik, dan akses internet. Kendala yang perlu diantisipasi adalah variasi kemampuan digital peserta, kesiapan data respons siswa, dan kemungkinan koneksi internet yang tidak stabil. Berdasarkan temuan tersebut, tim menyiapkan modul langkah demi langkah, contoh file data, paket soal pilihan ganda sebanyak 30 butir, serta data latihan cadangan. Penyiapan ini penting karena keberhasilan pelatihan berbasis teknologi sangat dipengaruhi oleh dukungan teknis, kesiapan pengguna, dan kemudahan alur penggunaan aplikasi (Alatli, 2020; Atabek, 2020; Gonzales et al., 2025; Shahzad et al., 2025; Theodorio, 2024). Oleh karena itu pelaksanaan pelatihan analisis butir berbasis teknologi memiliki dukungan fasilitas dasar yang cukup memadai, akan tetapi beberapa hal perlu diperhatikan yaitu mengantisipasi perbedaan kemampuan digital peserta, kesiapan data, dan stabilitas koneksi internet.

2. Tahap Pelaksanaan: Sosialisasi, Praktik 3S-CTT, dan Interpretasi Output

Tahap pelaksanaan diawali dengan sosialisasi konsep dasar analisis butir. Peserta diperkenalkan pada fungsi tingkat kesukaran, daya pembeda, reliabilitas, dan efektivitas pengecoh dalam menilai kualitas soal pilihan ganda. Penekanan diberikan pada hubungan antara data respons siswa dan keputusan perbaikan butir, sehingga guru memahami bahwa hasil analisis tidak hanya berupa angka, tetapi menjadi dasar untuk mempertahankan, merevisi, atau mengganti soal. Pemahaman ini sejalan dengan peran CTT sebagai pendekatan praktis untuk membaca kualitas butir dan mendukung keputusan evaluasi di kelas (Ayanwale et al., 2022; Hu et al., 2021; Rezigalla et al., 2024; Vicent & Shanmugam, 2020).

Setelah sosialisasi, peserta mengikuti demonstrasi penggunaan software 3S-CTT. Demonstrasi mencakup alur masuk ke aplikasi, format data, input respons siswa, pengaturan kunci jawaban, proses analisis, penyimpanan *output*, dan pembacaan hasil. Pada tahap praktik, peserta mencoba menjalankan analisis menggunakan data latihan atau data soal yang telah

disiapkan. Beberapa kendala teknis yang muncul antara lain ketidaksesuaian format data, kekeliruan input kunci jawaban, dan kebingungan awal dalam membaca indikator *output*. Kendala tersebut diatasi melalui pendampingan langsung, contoh ulang, dan pemeriksaan bersama terhadap file yang digunakan peserta, seperti terlihat pada Gambar 1.



Gambar 1. Pelaksanaan pelatihan analisis butir soal menggunakan software 3S-CTT.

Pada Gambar 1 memperlihatkan suasana pelatihan di ruang kelas mitra. Pemateri menjelaskan alur penggunaan 3S-CTT melalui tayangan presentasi, sedangkan peserta mengikuti penjelasan dan praktik menggunakan perangkat masing-masing. Dokumentasi ini menunjukkan bahwa kegiatan tidak hanya berbentuk ceramah, tetapi juga melibatkan praktik langsung, diskusi, dan pendampingan teknis selama proses analisis berlangsung.

Sesi interpretasi *output* menjadi bagian penting dalam pelaksanaan kegiatan. Peserta dibimbing membaca nilai tingkat kesukaran untuk menentukan apakah suatu butir terlalu mudah, sedang, atau terlalu sukar; membaca daya pembeda untuk melihat kemampuan butir dalam membedakan peserta didik berkemampuan tinggi dan rendah; serta memeriksa efektivitas pengecoh untuk mengetahui apakah pilihan jawaban yang salah berfungsi secara memadai. *Output* analisis kemudian digunakan untuk menyusun rekomendasi, seperti mempertahankan butir yang baik, merevisi butir yang bermasalah, mengganti pengecoh yang tidak dipilih, atau membuang butir yang tidak memenuhi kriteria. Dengan demikian, software berperan sebagai alat bantu, sedangkan keputusan akhir tetap memerlukan pertimbangan profesional guru sesuai konteks pembelajaran.

3. Tahap Evaluasi: Capaian, Kendala, dan Keberlanjutan Program

Tahap evaluasi dilakukan untuk melihat ketercapaian tujuan pelatihan. Evaluasi menggunakan *pre-test* dan *post-test* untuk mengukur pemahaman konsep, checklist keterampilan untuk menilai kemampuan teknis, telaah *output* untuk melihat keberhasilan peserta menjalankan analisis, serta diskusi reflektif untuk menggali kendala dan kebutuhan tindak lanjut. Berdasarkan proses pelatihan, peserta menunjukkan peningkatan

pemahaman terhadap fungsi analisis butir dan mulai mampu mengikuti alur teknis penggunaan 3S-CTT. Hasil praktik juga menunjukkan bahwa peserta dapat menghasilkan *output* analisis dan menggunakan *output* tersebut sebagai dasar awal penyusunan rekomendasi perbaikan soal.

Kendala utama yang muncul selama kegiatan adalah perbedaan kemampuan digital antarpeserta, ketelitian dalam menyiapkan format data, dan kebutuhan waktu tambahan untuk memahami makna statistik pada *output*. Hal ini menunjukkan bahwa pelatihan analisis butir sebaiknya tidak berhenti pada pengenalan aplikasi, tetapi perlu dilanjutkan dengan pendampingan interpretasi dan praktik berkala menggunakan data nyata dari mata pelajaran masing-masing. Temuan ini sejalan dengan kajian yang menyatakan bahwa pemilihan aplikasi analisis perlu disesuaikan dengan kemampuan pengguna, kompleksitas data, dan kebutuhan interpretasi hasil (Nurjanah et al., 2024; Susanto et al., 2026), seperti terlihat pada Tabel 3.

Tabel 3. Temuan Evaluasi dan Tindak Lanjut Kegiatan

Aspek Evaluasi	Teknik Evaluasi	Temuan Utama	Tindak Lanjut
Pemahaman konsep	<i>Pre-test</i> , <i>post-test</i> , dan diskusi	Peserta memahami fungsi dasar CTT, terutama tingkat kesukaran, daya pembeda, reliabilitas, dan pengecoh.	Memberikan contoh interpretasi yang lebih banyak dari berbagai mata pelajaran.
Keterampilan teknis	Observasi dan checklist	Peserta dapat mengikuti alur input data, pengaturan kunci, proses analisis, dan penyimpanan <i>output</i> dengan pendampingan.	Menyediakan panduan singkat dan template data yang mudah digunakan.
<i>Output</i> praktik	Telaah file <i>output</i> /laporan praktik	Peserta menghasilkan <i>output</i> analisis butir dari data latihan atau data yang disiapkan.	Mendorong penggunaan data nyata dari penilaian semester berikutnya.
Interpretasi hasil	Rubrik interpretasi dan tanya jawab	Peserta mulai mampu menyusun rekomendasi perbaikan butir berdasarkan indikator <i>output</i> .	Melakukan klinik daring untuk mendampingi interpretasi lanjutan.

Keberlanjutan program diarahkan melalui konsultasi daring antara tim pelaksana dan guru peserta. Jalur pendampingan ini diperlukan agar guru memperoleh bantuan ketika menggunakan 3S-CTT pada data penilaian yang sesungguhnya. Selain itu, sekolah mitra dapat mengembangkan arsip hasil analisis butir sebagai dasar penyusunan bank soal dan perbaikan instrumen evaluasi secara berkala. Dengan langkah tersebut, kegiatan pelatihan tidak hanya berdampak pada peningkatan keterampilan individu, tetapi juga

mendukung terbentuknya budaya evaluasi pembelajaran berbasis data di sekolah.

D. SIMPULAN DAN SARAN

Kegiatan pelatihan analisis butir soal menggunakan software 3S-CTT menjadi upaya penguatan literasi asesmen guru dalam melakukan evaluasi pembelajaran berbasis data. Revisi program disusun melalui tiga tahap utama, yaitu pra-kegiatan, pelaksanaan, dan evaluasi. Tahap pra-kegiatan menghasilkan identifikasi kebutuhan guru, modul pelatihan, dan data praktik. Tahap pelaksanaan membantu peserta memahami konsep CTT, menjalankan analisis butir, serta membaca *output* tingkat kesukaran, daya pembeda, reliabilitas, dan efektivitas pengecoh. Tahap evaluasi menunjukkan bahwa peserta mulai mampu menghasilkan *output* analisis dan menyusun rekomendasi perbaikan butir dengan pendampingan. Dengan demikian, pelatihan 3S-CTT berpotensi menjadi model praktis untuk memperkuat budaya asesmen berbasis bukti di sekolah mitra.

Saran untuk kegiatan berikutnya adalah memperluas jumlah peserta, menggunakan data nyata dari berbagai mata pelajaran, menambah sesi khusus interpretasi *output*, dan menyediakan pendampingan daring setelah pelatihan. Sekolah mitra juga disarankan mengarsipkan hasil analisis butir sebagai dasar penyusunan bank soal dan perbaikan instrumen evaluasi pada periode pembelajaran berikutnya. Jumlah butir soal yang digunakan dalam naskah ini perlu disesuaikan dengan data aktual kegiatan apabila berbeda dari data latihan yang dicantumkan.

UCAPAN TERIMA KASIH

Ucapan terima kasih disampaikan kepada Universitas Muhammadiyah Mataram melalui skema Pengabdian kepada Masyarakat Internal UMMAT Tahun 2026 atas dukungan pendanaan dan fasilitasi kegiatan. Terima kasih juga disampaikan kepada sekolah mitra, kepala sekolah, guru peserta, serta mahasiswa yang membantu aspek administrasi, dokumentasi, dan pendampingan teknis selama kegiatan berlangsung.

DAFTAR RUJUKAN

- Alatli, B. (2020). A Systematic Review of Research Articles on Measurement Invariance in Education and Psychology. *International Journal of Assessment Tools in Education*, 7(4), 607–630. <https://doi.org/10.21449/ijate.738560>
- Atabek, O. (2020). Experienced educators' suggestions for solutions to the challenges to technology integration. *Education and Information Technologies*, 25(6), 5669–5685. <https://doi.org/https://doi.org/10.1007/s10639-020-10243-y>
- Ayanwale, M. A., Masopha, julia chere, & Morena, M. (2022). The Classical Test or Item Response Measurement Theory: The Status of the Framework at the Examination Council of. *International Journal of Learning, Teaching and Educational Research*, 21(8), 384–406.
- Choi, Y., & Asilkalkan, A. (2019). R Packages for Item Response Theory Analysis :

- Descriptions and Features R Packages for Item Response Theory Analysis : Descriptions and. *Measurement: Interdisciplinary Research and Perspectives*, 17(3), 168–175. <https://doi.org/10.1080/15366367.2019.1586404>
- Gonzales, P. S., Diz, J. S., Iglesias, D., & Rio, J. F. (2025). Exploring physical education teachers' willingness and barriers to integrating digital technology in their lessons. *Education and Information Technologies*, 30(5), 5965–5987. <https://doi.org/10.1007/s10639-024-13060-9>
- Hayat, B. (2021). Klasika : Program Analisis Item dan Tes dengan Pendekatan Klasik. *JP3I (Jurnal Psikologi Dan Pendidikan Indonesia)*, 10(1), 1–11. <https://doi.org/http://dx.doi.org/10.15408/jp3i.v10i1.20551>
- Hu, Z., Lin, L., Wang, Y., & Li, J. (2021). The Integration of Classical Testing Theory and Item Response Theory. *Psychology*, 12, 1397–1409. <https://doi.org/10.4236/psych.2021.129088>
- Nurjanah, S., Iqbal, M., Zafrullah, Z., Mahmud, M. N., Seran, D. S. F., Suardi, I. K., & Arriza, L. (2024). Psychometric quality of multiple-choice tests under classical test theory (CTT): AnBuso, Iteman, and R. *Jurnal Penelitian Dan Evaluasi Pendidikan*, 28(2), 161–172.
- Qadir, A., Huda, N., & Hermina, D. (2024). Analisis Butir Tes : Tingkat Kesukaran, Daya Pembeda Dan Efektivitas Pengecoh. *Al Furqan : Jurnal Agama, Sosial, Dan Budaya*, 3(3), 1450–1467.
- Rezigalla, A. A., Mohammed, A., Seid, E., Eleragi, A., Elhussein, A. B., Alfaifi, J., Alghamdi, M. A., Ameer, A. Y. Al, Ibrahim, A., Yahia, O., Mohammed, O. A., Ishag, M., & Adam, E. (2024). Item analysis : the impact of distractor efficiency on the difficulty index and discrimination power of multiple-choice items. *BMC Medical Education*, 24(445), 1–7.
- Ridha, A. R., Luthfi, M., & Mudijono, M. (2025). Evaluasi Butir Soal Secara Klasik Dalam Evaluasi Pendidikan. *JIPDAS : Jurnal Ilmiah Pendidikan Dasar*, 4(1), 100–103.
- Rohmatdi, A., Martono, M., Zafrullah, Z., & Safitri, R. (2024). Optimizing Question Quality in Junior High School Exams : Classical Test Theory Evaluation with ITEMAN 4 . 3. *Al-Ishlah: Jurnal Pendidikan*, 16(4), 4363–4375. <https://doi.org/10.35445/alishlah.v16i4.5671>
- Shahzad, M. F., Xu, S., An, X., & Asif, M. (2025). Are Generative AI Technologies Transforming Education for the 21st Century ? Research Trends , Challenges , and Benefits. *Sage Open*, 15(3), 1–18. <https://doi.org/10.1177/21582440251368594>
- Susanto, H. P., Abadi, A. M., & Retnawati, H. (2025). Development of irtawsi : A User-Friendly R Package for IRT Analysis. *Jurnal Pengukuran Psikologi Dan Pendidikan Indonesia (JP3I)*, 14(1), 1–23.
- Susanto, H. P., Fathurrahman, Purnamasari, M. I., Qudsiyah, K., Sabaryati, J., Darmayasa, jero B., Suciati, Prihono, E. W., & Kusuma, M. (2026). Pengembangan Software Analisis Butir Klasik Untuk Calon Guru. *Jurnal Pendidikan*, 14(01), 7–21.
- Susanto, H. P., Qudsiyah, K., Meifiani, N. I., Apriyani, D. C. N., Hidayat, T., & Purnamasari, M. I. (2022). Sharing Pengetahuan Analisis Butir Tes dan Praktiknya Menggunakan Program R. *Journal of Social Empowerment*, 07(02), 119–127. <https://doi.org/10.21137/jse.2022.7.2.5>
- Theodorio, A. O. (2024). Examining the support required by educators for successful technology integration in teacher professional development program. *Cogent Education*, 11(1), 1–11. <https://doi.org/10.1080/2331186X.2023.2298607>
- Vicent, W., & Shanmugam, S. K. S. (2020). The Role of Classical Test Theory to Determine the Quality of Classroom Peran Teori Uji Klasik untuk Menentukan Kualitas Item Tes Pengajaran di Kelas. *PEDAGOGIA: Jurnal Pendidikan*, 9(1), 5–34.